

IN

NCM105 ORF DNA Sequence [1 to 1116] -> 1-phase Translation

DNA sequence 1116 b.p. ATGGCTGCACC ... AGTCGAACTGA linear

complete new ETS ORF region DNA sequence (confirmed on 6/14/96).

37/ amino acids

$M_w = 41428 \text{ Dalton}$

[illegible]

Fig. 1

Fig. 2a

MAATCEISNIFSNYFSAMYSSDSTLASVPPAATFGADDLVLTLSNPQMSLEG	53
TEKASWLGEQ	106
<u>PQFWSKTQVLDWISYQVEKNKYDASAIDFSRCDMDGATLCNCA</u>	
LEELRLVFGPLGDQLHAQLRDLTSSSSDELSWIIELLEKDGMAFQEAALDPGPF	159
DQGSPPAQELLDDGQQASPYHPGSCGAG	212
<u>APSPGSSDVSTAGTGASRSSHSSDS</u>	
<u>GGSDVDLDPTDGKLLFPDGFDRDCKKG</u>	265
SKHAPRGTHLWEEFIRDILIHPELNEGLMKWENRHEGVFKFLRSEAVAQLWGQK	318
<u>KKNSNMTYEKLSRAMRYYYKREILERVDGRRLLVYKF</u>	371
<u>GKNSSGWKEEEVLQSRN</u>	

Fig. 2b

Consensus⁺: P W V W W E G LC
 ESX (64-103): PQFWSKTQVLWDWISYQVEKNKYDASAIDEFRCMDGATLC
 P+ W++T V DW+ + V N++ +DF + M+GA LC
 ETS-1 (69-106): PRQWTEHVDRDWMWAV--NEFSLKGVDFQKFCMNGAALC

Fig. 2c

ESX (188-238): APSPGSSDVSTAGTGASRSSHSSDGGSDVDLDPTDGKLFPSDGFRCCKG
 APS S S++ + S SS SS S SD + + L PS F G
 SOX4 (370-420): APSSAPSHASSASSHSSSSSSGSSDDEFEDDLLDNPSSNFESMSLG

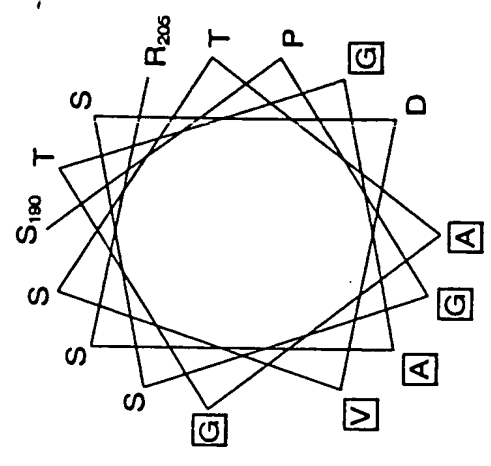


Fig. 2d

Consensus⁺: α_1 β_1 β_2 α_2 "turn" α_3 β_3 "wing" β_4
 LWQFLL LL D I W FK VAR WG K P MNY KLSR LRYYY I K G R Y F
 * * * * *
 ESX (274-354): HLWEFIRDILIHPELNEGLMKWENRHEGVFKFLRSEAVAQLWGQKKNSNMTYEKLSRAMRYYYKREILERVDRRLVYKF
 +LWEF+ +L +KW R +G+FK + S+AV++LWG+ K +M YE + RA+RYYY+R IL +V+G+RLVY+F
 ELF-1 (209-289): YLWEFLLALLQDKATCPKYIKWTQREKGIFKLVDKAVSRVLWGKHKNKPDNMNYETMGRALRYYYQRGILAKVEGQRLVYQF

Human ESX Protein Sequence

MAATCEISNIFSNYFSAMYSSDSTLASVPPAATFGADDLVLTLSNPQMSLEG 53
 TEKASWLGEQPFWSKTQVLDWISYQVEKNKYDASAIQFSRCDMDGATLCNCA 106
 LEELRLVFGPLGDLHAQLRDLTSSSSDELSWYIELLQGGMAFQALDPGPF 159
 DQGSPFAQELLDDGQQAAPYHPGSCGAGAPSGSSDVSTAGTGASRSSHSDS 212
 GGSDVDLDPDTDGKLFPSDGFRCCKGDPKHGKRKRGRPRKLSKEYWOCLEGKK 265
 SKHAPRGTHLWEIFRDLIHPELNEGLMKWENRHEGVFKFLRSEAVAQLWGQK 318
 KKNSNMTYEKLSRAMRYYYKREILERVDRRLVYKFGKNSSGWKEEVLQSRN 371

Fig. 2e

mESX Genomic Organization and Gene Product

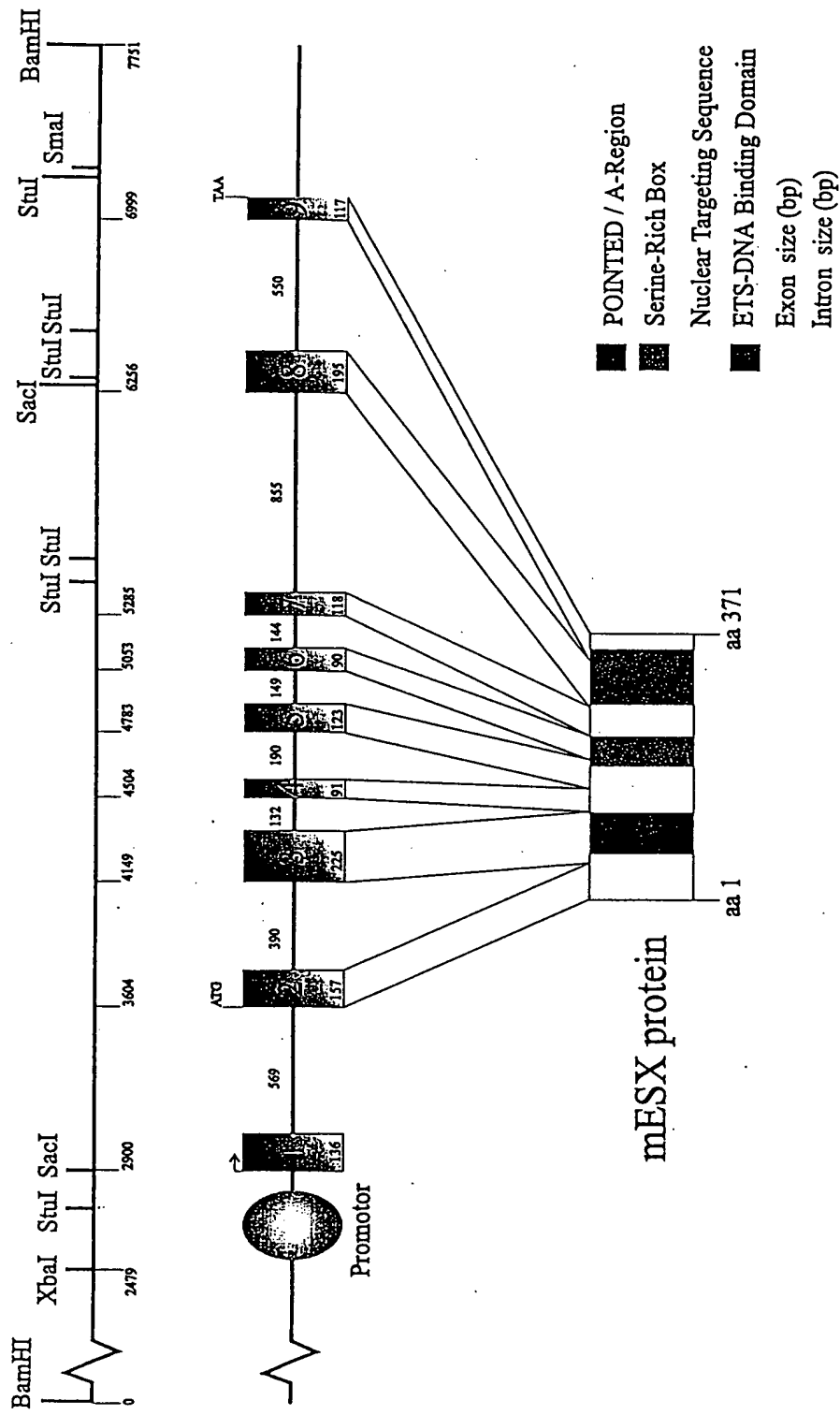


Fig. 3

Human ESX Exon/Intron Junctions

1/1	ATG GCT GCA ACC TGT GAG ATT AGC AAC ATT	31/11	TTT AGC AAC TAC TTC AGT GCG ATG TAC AGC
Met ala ala thr cys glu ile ser asn ile		phe ser asn tyr phe ser ala met tyr ser	
61/21	TCG GAG GAC TCC ACC CTG GCC TCT GTT CCC	91/31	CCT GCT GCC ACC TTT GGG GCC GAT GAC TTG
ser glu asp ser thr leu ala ser val pro		pro ala ala thr phe gly ala asp asp leu	
121/41	GTA CTG ACC CTG AGC AAC CCC CAG ATG TCA	151/51	TTG GAG GGT ACA GAG AAG GCC AGC TGG TTG
val leu thr leu ser asn pro gln met ser		leu glu gly thr glu lys ala ser trp leu	
181/61	GGG GAA CAG CCC CAG TTC TGG TCG AAG ACG	211/71	CAG GTT CTG GAC TGG ATC AGC TAC CAA GTG
gly glu gln pro gln phe trp ser lys thr		gln val leu asp trp ile ser tyr gln val	
241/81	GAG AAG AAC AAG TAC GAC GCA AGC GCC ATT	271/91	GAC TTC TCA CGA TGT GAC ATG GAT GGC GCC
glu lys asn lys tyr asp ala ser ala ile		asp phe ser arg cys asp met asp gly ala	
301/101	ACC CTC TGC AAT TGT GCC CTT GAG GAG CTG	331/111	CGT CTG GTC TTT GGG CCT CTG GGG GAC CAA
thr leu cys asn cys ala leu glu glu leu		arg leu val phe gly pro leu gly asp gln	
361/121	CTC CAT GCC CAG CTG CGA GAC CTC ACT TCC	391/131	AGC TCT TCT GAT GAG CTC AGT TGG ATC ATT
leu his ala gln leu arg asp leu thr ser	E3▼E4	ser ser ser asp glu leu ser trp ile ile	
421/141	GAG CTG CTG GAG AAG GAT GGC ATG GCC TTC	451/151	CAG GAG GCC CTA GAC CCA GGG CCC TTT GAC
glu leu leu glu lys asp gly met ala phe		gln glu ala leu asp pro gly pro phe asp	E4▼E5
481/161	CAG GGC AGC CCC TTT GCC CAG GAG CTG CTG	511/171	GAC GAC GGT CAG CAA GCC AGC CCC TAC CAC
gln gly ser pro phe ala gln glu leu leu		asp asp gly gln gln ala ser pro tyr his	
541/181	CCC GGC AGC TGT GGC GCA GGA GCC CCG TCC	571/191	CCT GGC AGC TCT GAC GTC TCC ACC GCA GGC
pro gly ser cys gly ala gly ala pro ser		pro gly ser ser asp val ser thr ala gly	E5▼E6
601/201	ACT GGT GCT TCT CGG AGC TCC CAC TCC TCA	631/211	GAC TCC GGT GGA AGT GAC GTG GAC CTG GAT
thr gly ala ser arg ser ser his ser ser		asp ser gly gly ser asp val asp leu asp	
661/221	CCC ACT GAT GGC AAG CTC TTC CCC AGC GAT	691/231	GGT TTT CGT GAC TGC AAG AAG GGG GAT CCC
pro thr asp gly lys leu phe pro ser asp	E6▼E7	gly phe arg asp cys lys lys gly asp pro	
721/241	AAG CAC GGC AAG CGG AAA CGA GGC CGG CCC	751/251	CGA AAG CTG AGC AAA GAG TAC TGG GAC TGT
lys his gly lys arg lys arg gly arg pro		arg lys leu ser lys glu tyr trp asp cys	
781/261	CTC GAG GGC AAG AAG AGC AAG CAC GCG CCC	811/271	AGA GGC ACC CAC CTG TGG GAG TTC ATC CCG
leu glu gly lys lys ser lys his ala pro	E7▼E8	arg gly thr his leu trp glu phe ile arg	
841/281	GAC ATC CTC ATC CAC CCG GAG CTC AAC GAG	871/291	GGC CTC ATG AAG TGG GAG AAT CCG CAT GAA
asp ile leu ile his pro glu leu asn glu		gly leu met lys trp glu asn arg his glu	
901/301	GGC GTC TTC aag TTC CTG CGC TCC GAG GCT	931/311	GTG GCC CAA CTA TGG GGC CAA AAG AAA AAG
gly val phe lys phe leu arg ser glu ala		val ala gln leu trp gly gln lys lys lys	
961/321	AAC AGC AAC ATG ACC TAC GAG AAG CTG AGC	991/331	CGG GCC ATG AGG TAC TAC TAC AAA CGG GAG
asn ser asn met thr tyr glu lys leu ser	E8▼E9	arg ala met arg tyr tyr tyr lys arg glu	
1021/341	ATC CTG GAA CGG GTG GAT GGC CGG CGA CTC	1051/351	GTC TAC AAG TTT GGC AAA AAC TCA AGC GGC
ile leu glu arg val asp gly arg arg leu		val tyr lys phe gly lys asn ser ser gly	
1081/361	TGG AAG GAG GAA GAG GTT CTC CAG AGT CGG	1111/371	AAC TGA
trp lys glu glu glu val leu gln ser arg		asn OPA	

Fig. 4

ESX Primary Structure and Domain Homologies.

mESX 1 MAATCEISNVFSNYFNAMYSSDPTLAPAPP.TTFTGEDLVLTLLNNQOMT 49
hESX 1 MAATCEISNIFSNYFSAMYSSDSTLASVPPAATFGADDLVLTLSNPQMS 50

E_2/E_1
↓

mESX 50 LEGPEKASWTSER [REDACTED] 99
hESX 51 LEGTEKASWLGEQ [REDACTED] 100

E_3/E_1
↓

mESX 100 [REDACTED] TSNSSDELSWIIELLEKDGMSF 149
hESX 101 [REDACTED] TSSSSDELSWIIELLEKDGMAF 150

E_1/E_1 ↓ E_1/E_1 ↓

mESX 150 QESLGLDLPDQSGSPFAQELLDDGRQASPYCSTYGPAPSPGSSDVSTA 199
hESX 151 QEAL.DPGPFDQSGSPFAQELLDDGQQASPYHPGSCGAGAPSPGSSDVSTA 199

E_1/E_1 ↓

mESX 200 [REDACTED] DDFTDYKKGEPPKHGKRKRGR 249
hESX 200 [REDACTED] DGFRDCKKGDPPKHGKRKRGR 249

E_1/E_1 ↓

mESX 250 PRKLSKEYWDCLEGKKS KHAPRGT [REDACTED] 299
hESX 250 PRKLSKEYWDCLEGKKS KHAPRGT [REDACTED] 299

E_1/E_1 ↓

mESX 300 [REDACTED] 349
hESX 300 [REDACTED] 349

mESX 350 [REDACTED] GKNSSGWKEEEVGESRN 371
hESX 350 [REDACTED] GKNSSGWKEEEVLQSRN 371

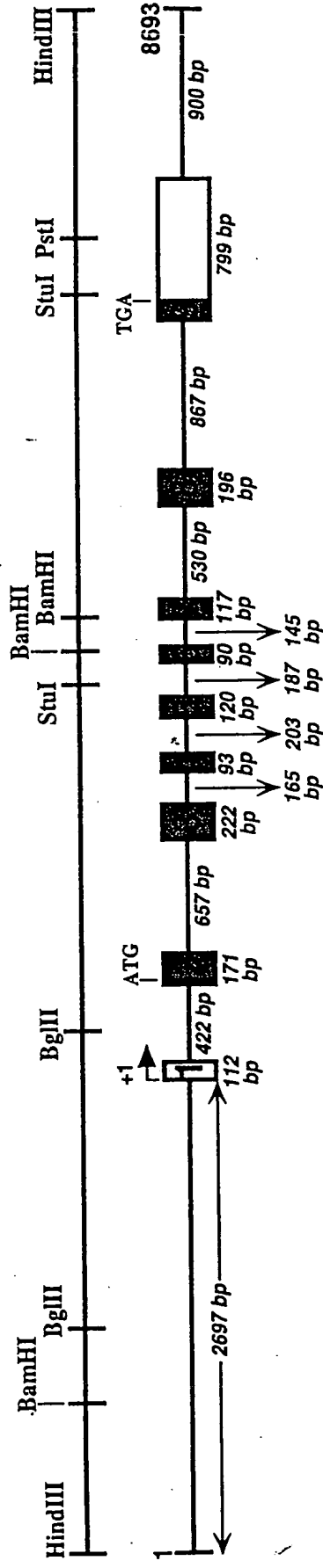
■ POINTED / A-Region
■ Serine-Rich Box
[] Nuclear Targeting Sequence
■ ETS-DNA Binding Domain

Figure 1 illustrates the experimental setup. A participant is seated at a table, looking at a video screen. The screen shows a target (a small circle) and a starting point (a larger circle). The participant's hand is positioned at the starting point. The video screen is connected to a computer system. The computer system is labeled 'Computer' and 'Video Screen'. The participant is labeled 'Participant'.

■	Ets	■	USF
■	AP-2	■	Oct
■	extended palindrome	■	NF-κB
■	SP1/GC box	□ □	CCAAT box & Inr element

1

Human ESX Genomic DNA Structure



Mouse ESX Genomic DNA Structure

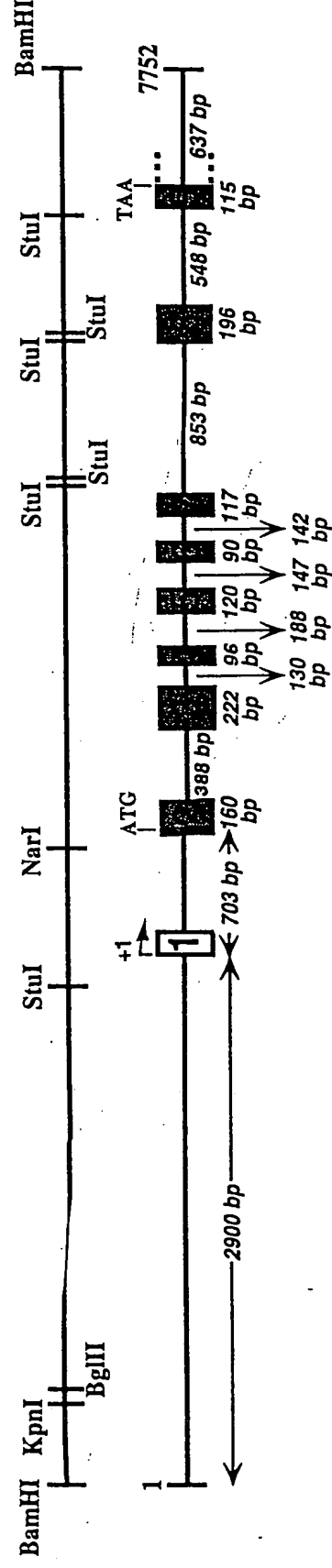
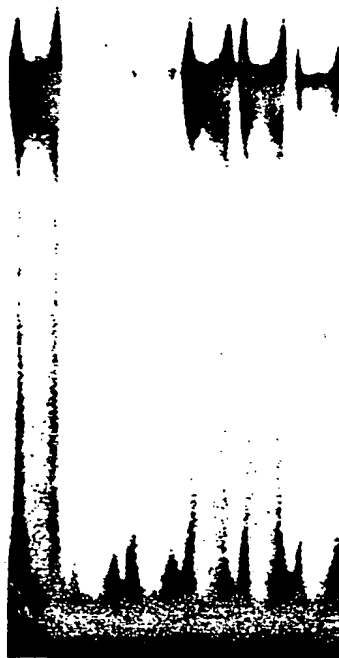


Fig. 7

Fig. 8a

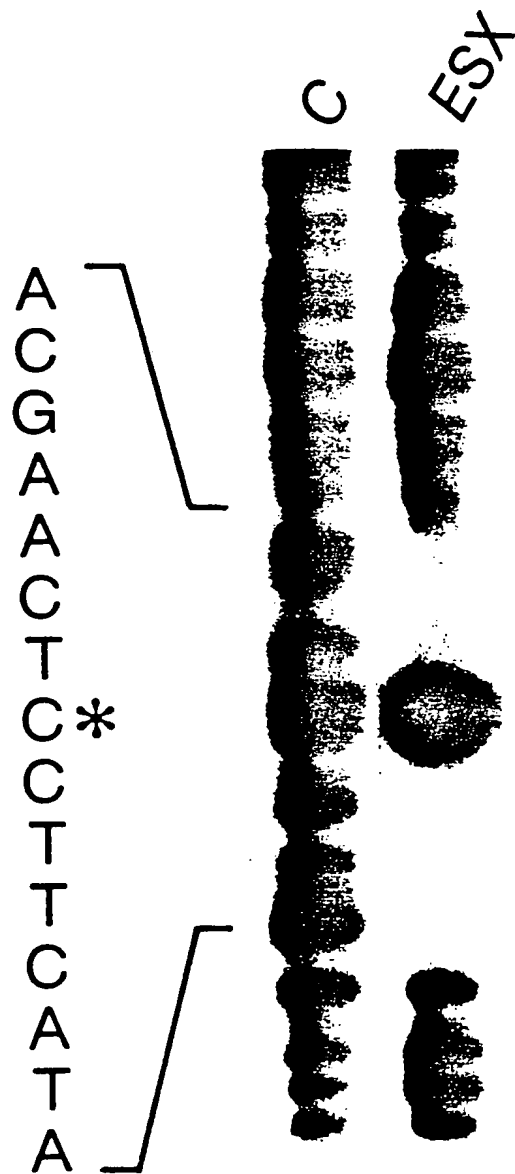
m1 m2 m3 m4 m5



WT 5' GGAGGAGGGCTGCTTGAGGAAGTATAAGAAT 3'
 m1 5' -----TA----- 3'
 m2 5' -----C----- 3'
 m3 5' -----AG----- 3'
 m4 5' -----CC----- 3'
 m5 5' -----C----- 3'

08978217.11259

Fig. 8b



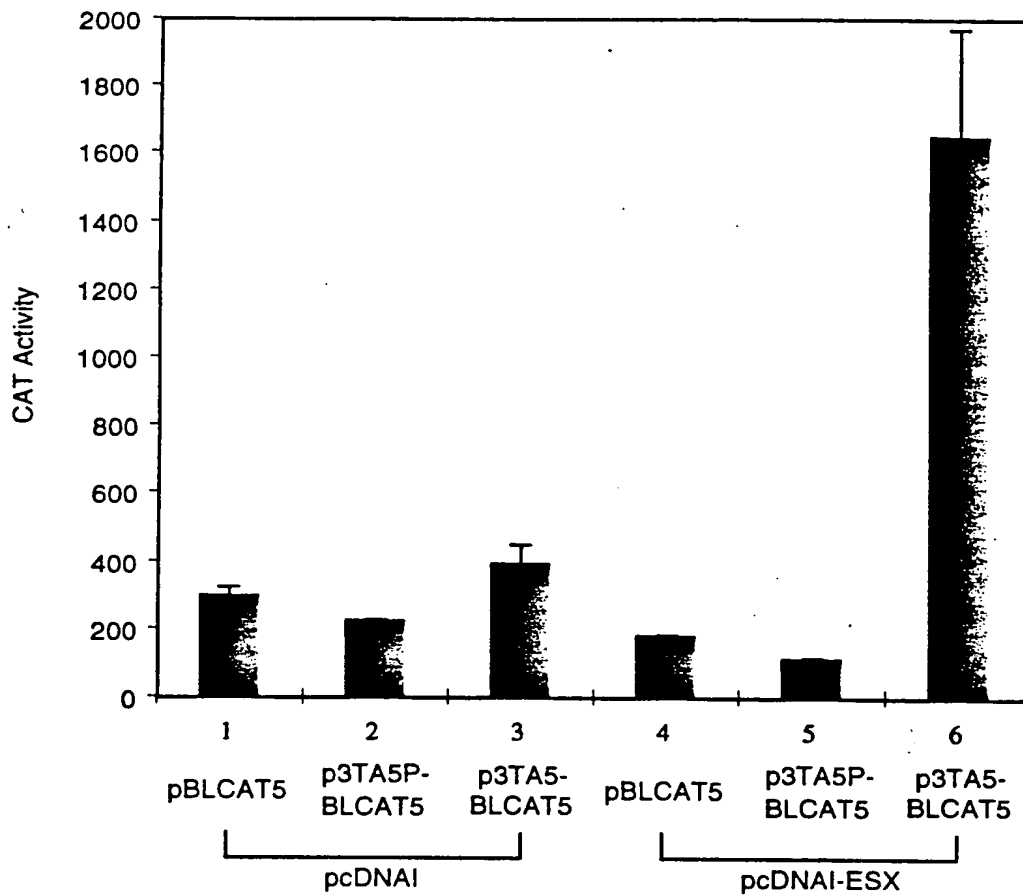


Fig. 8c

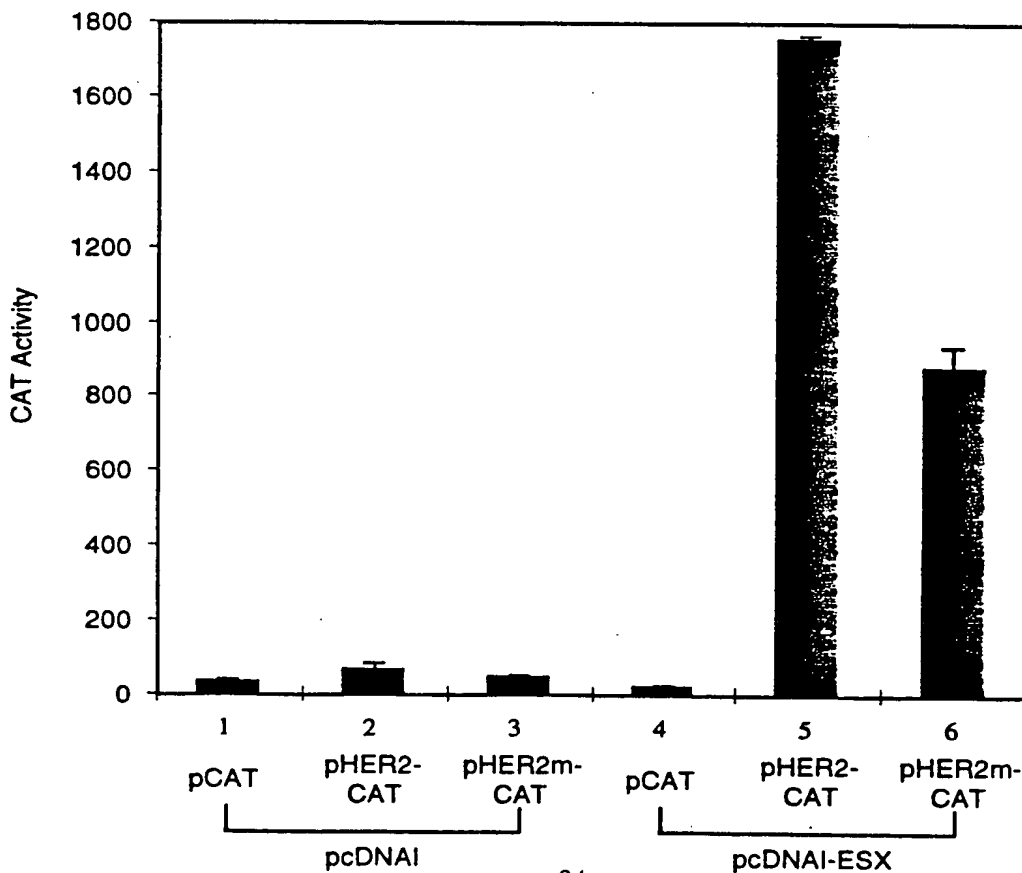
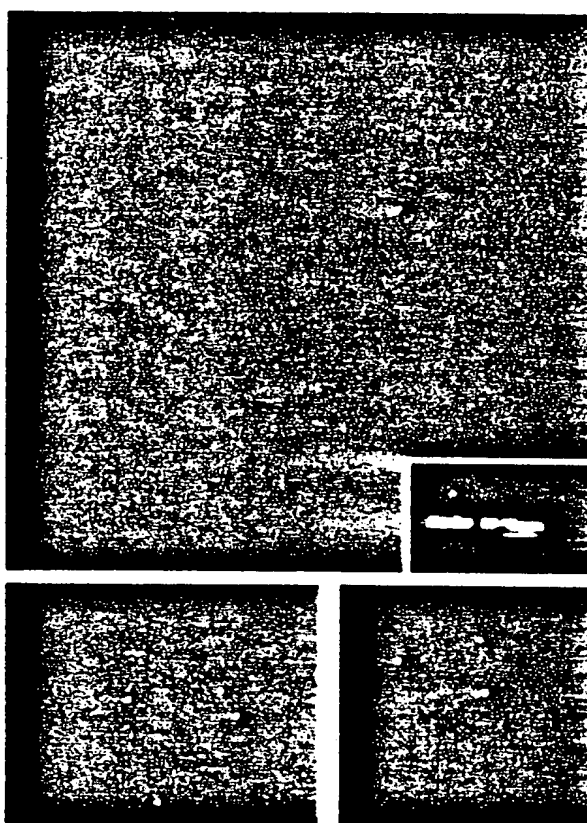


Fig. 8d

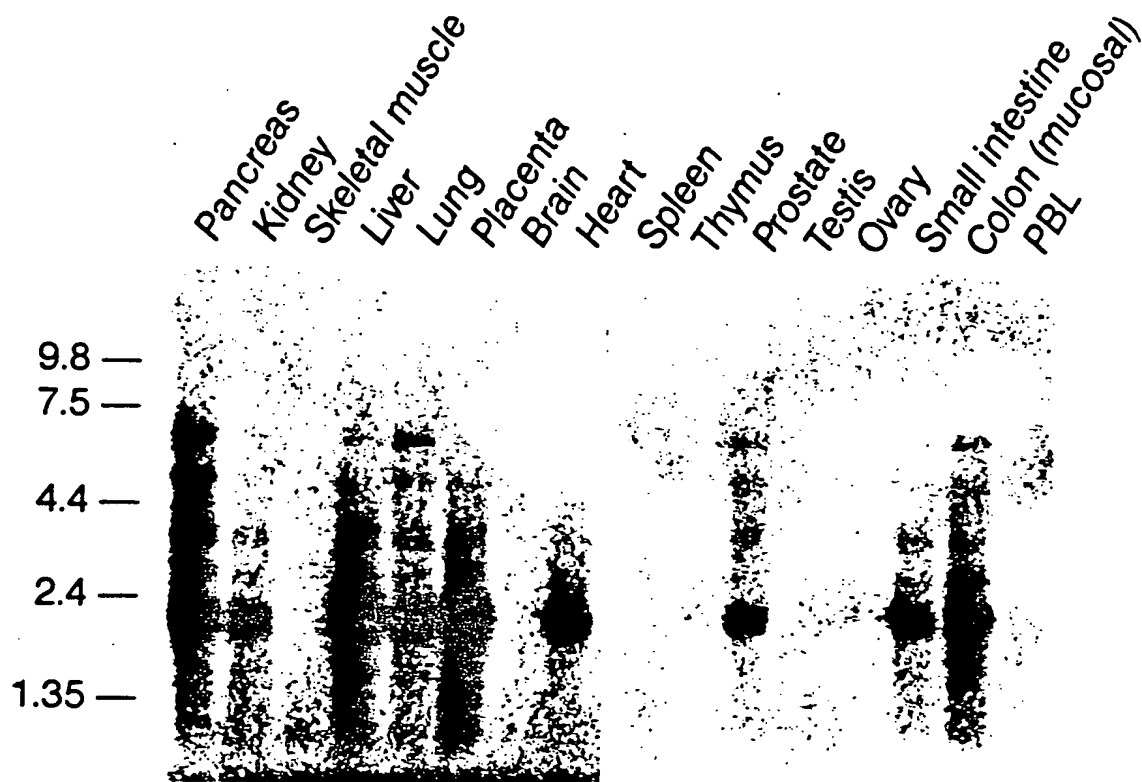
Fig. 8e



08978217 112597

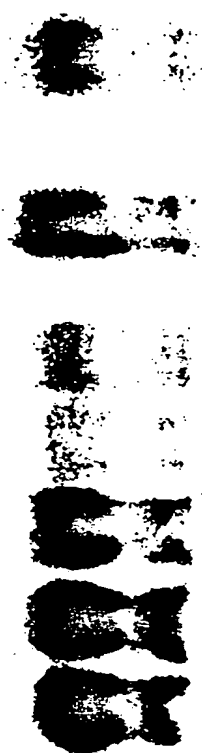
08978217 112597

Fig. 9a



HMEC
HBL-100
MCF-10A
MCF-7
MDA-231
BT-474
SK-BR-3
ZR-75-1

4.1 kb —
2.2 kb —



HER2- HER2+

Fig. 9b

08978217 112597

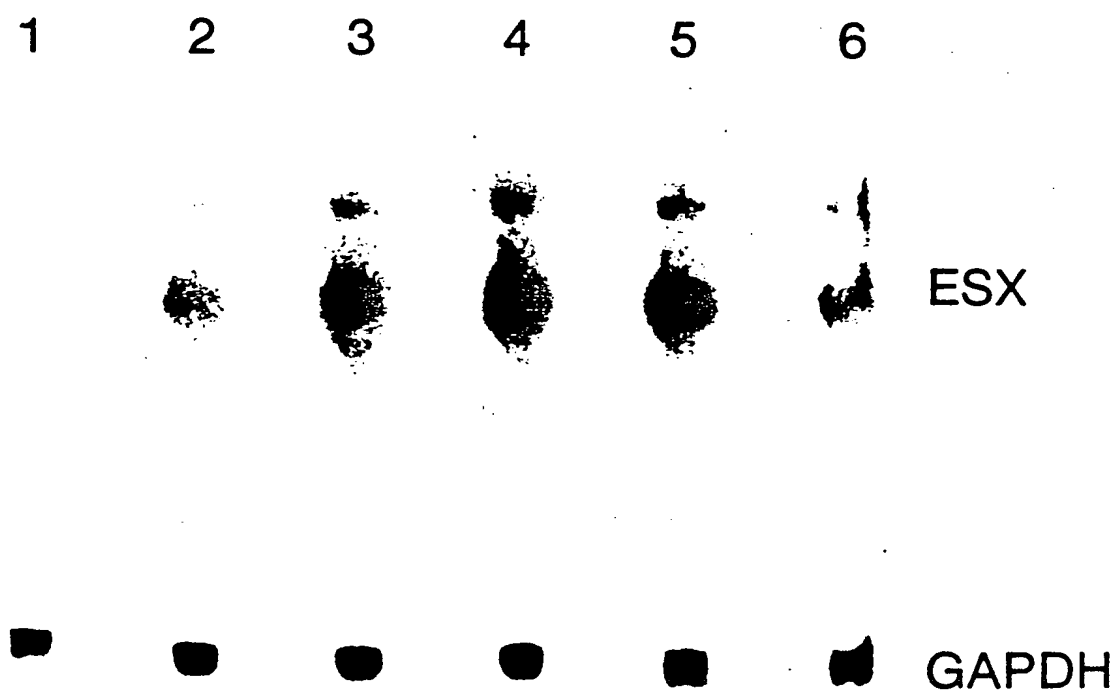


Fig. 9c

08978217 112597

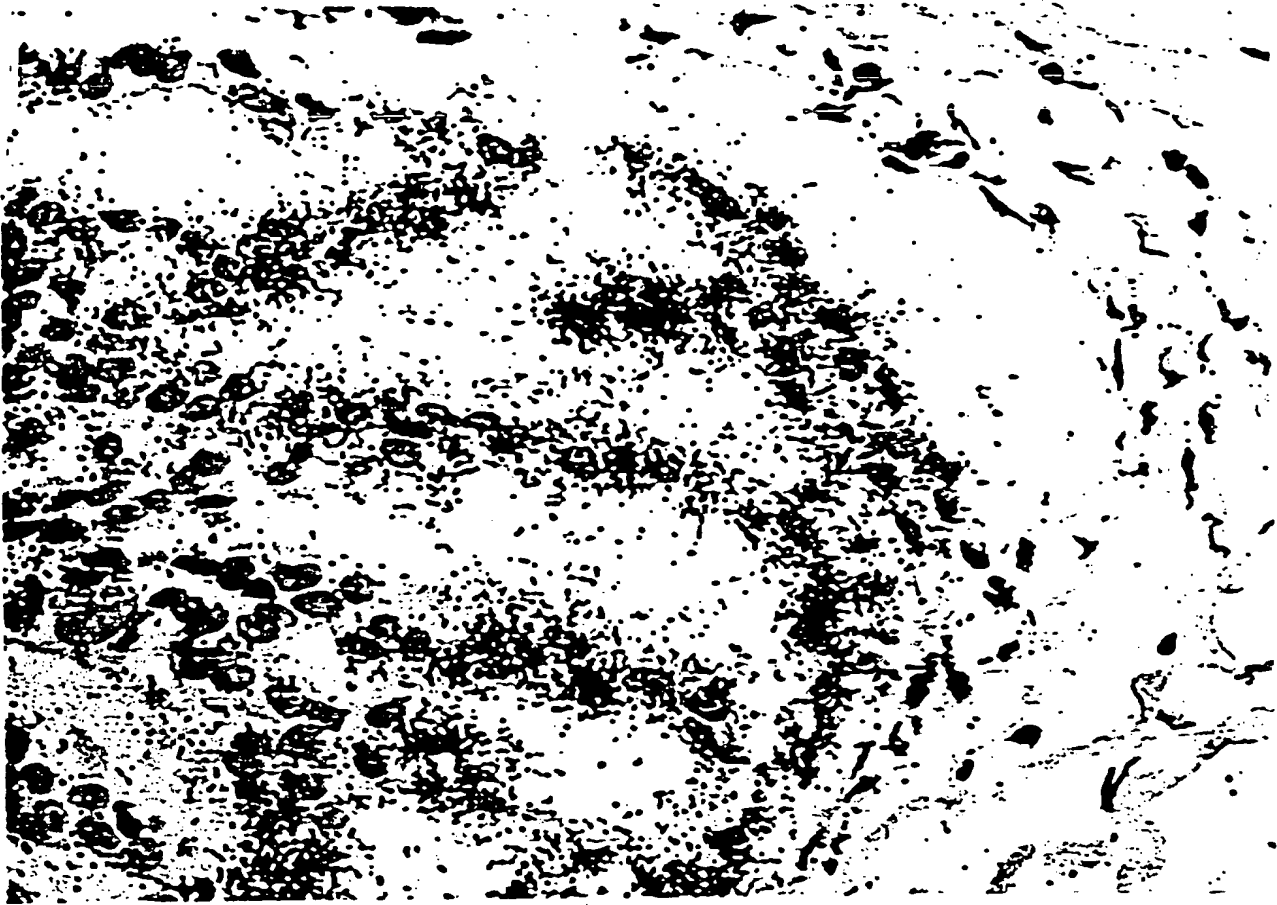


Fig. 10a

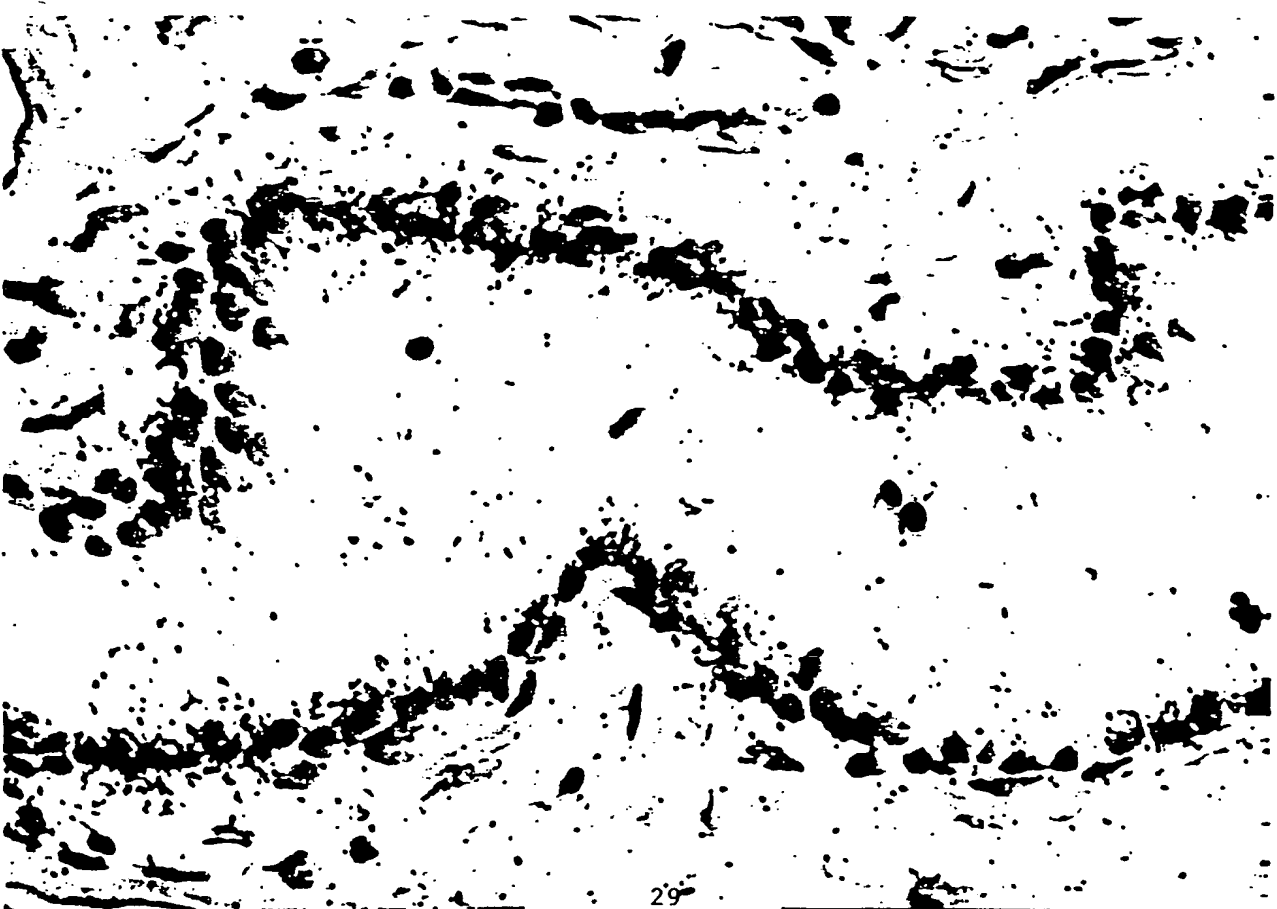


Fig. 10b

Mapping of hESX Activation Domain

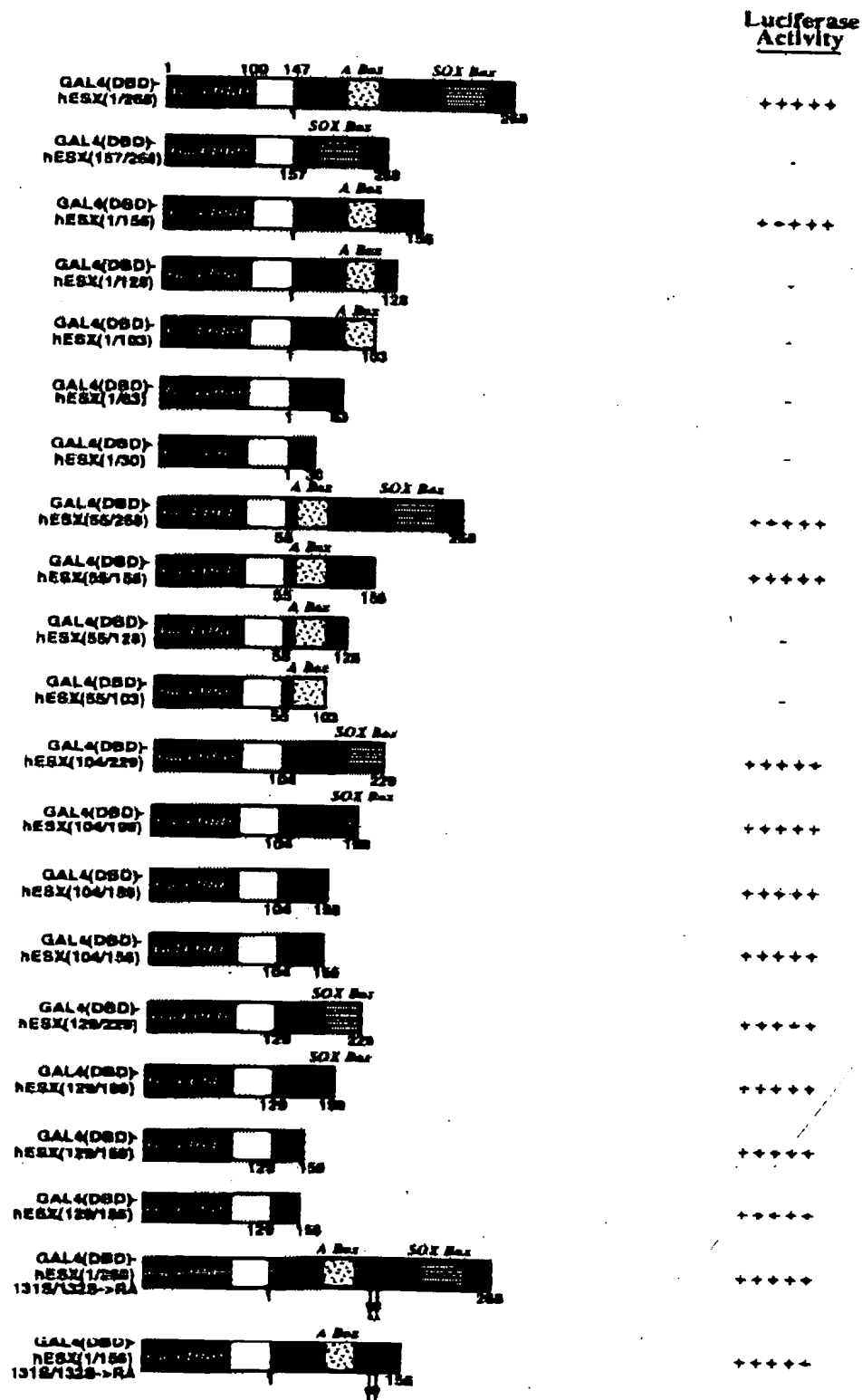


Fig. 11

08978247.11597